AD-A093 816 PITTSBURGH UNIV PA INST FOR STATISTICS AND APPLICATIONS F/G 12/1
CONCORDANT AND DISCORDANT MONOTONE CORRELATIONS AND THEIR EVALU--ETC(U)
OCT 80 G KIMELDORF, J H MAY, A R SAMPSON F49620-79-C-0161
UNCLASSIFIED TR-80-20 AFOSR-TR-80-1344 NL

1 OF 1
AD
A093816

END
DATE
FILMED
2 81
DTIC

AFOSR-TR- 80 - 1344

LEVEL

CONCORDANT AND DISCORDANT MONOTONE CORRELATIONS AND THEIR EVALUATION BY NONLINEAR OPTIMIZATION

George Kimeldorf[1]

Jerrold H. May

Allan R. Sampson[2]

University of Texas at Dallas

University of Pittsburgh

University of Pittsburgh

AD A093816

OCTOBER 1980

Technical Report No. 80-20

Institute for Statistics and Applications
Department of Mathematics and Statistics
University of Pittsburgh
Pittsburgh, PA. 15260

[1]The work of this author was supported by the National Science Foundation under Grant MCS-80-02152.

[2]The work of this author is sponsored by the Air Force Office of Scientific Research, Air Force Systems Command under Contract F49620-79-C-0161.

CONCORDANT AND DISCORDANT MONOTONE CORRELATIONS AND THEIR EVALUATION BY NONLINEAR OPTIMIZATION

by

George Kimeldorf, Jerrold H. May, and Allan R. Sampson

University of Texas at Dallas, University of Pittsburgh,

and University of Pittsburgh

ABSTRACT

This paper presents four new statistical measures of monotone relationship derived from the concept of monotone correlation. A nonlinear optimization algorithm is employed to evaluate these new measures, as well as the monotone correlation, for ordinal contingency tables. A computer program to implement the algorithm is developed, and is applied to several insightful examples to provide further understanding of the usefulness of these measures.

Key Words and Phrases: Monotone correlation, concordant monotone correlation, discordant monotone correlation, isoscaling, ordinal contingency tables, nonlinear optimization.

B

DISCLAIMER NOTICE

CONCORDANT AND DISCORDANT MONOTONE CORRELATIONS AND

THEIR EVALUATION BY NONLINEAR OPTIMIZATION

by

George Kimeldorf[1], Jerrold H. May, and Allan R. Sampson[2]

University of Texas at Dallas, University of Pittsburgh,

and University of Pittsburgh

1. Introduction and Statistical Background

Measuring and understanding the basis for the association between two random variables X and Y is extremely important for the intelligent application of statistics, as well as for more insight into the underlying bivariate probabilistic structures. The focus of this paper is on association between ordinal random variables, that is, random variables where the observed values have a natural ordering without necessarily having naturally ascribed numerical values. For example, the values may arise from questionnaire responses based on the five-point scale: strongly disagree, disagree, no opinion, agree, strongly agree. In measuring association between two ordinal variables using a measure based on assigning numerical values to the possible data values, it is natural to require that the resultant numerical measure of association not depend on the actual numerical values but only the orderings. This property can be described as monotone scale invariance. When Pearson's correlation coefficient is used by

[1]The work of this author was supported by the National Science Foundation under Grant MCS-80-02152.

[2]The work of this author is sponsored by the Air Force Office of Scientific Research, Air Force Systems Command under Contract F49620-79-C-0161.

assigning the values 1, ..., N to the scale levels, the resulting measure is not monotone scale invariant. For the five-point scale example, assigning 1 to strongly disagree, ..., 5 to strongly agree and then computing the Pearson correlation would not provide a monotone invariant measure of association.

One monotone invariant scale measure is the sup correlation ρ' introduced by Gebelein (1941), developed further by Sarmanov (1958a,b), Rényi (1959), and Lancaster (1969), and defined by $\rho'(X,Y) = \sup \rho(f(X), g(Y))$, where the supremum is taken over all Borel-measurable functions f,g such that $0 < \operatorname{Var} f(X) < \infty$ and $0 < \operatorname{Var} g(Y) < \infty$, where ρ is the Pearson correlation coefficient. For random variables (X,Y) jointly taking values on a finite rectangular lattice, there are computational methods for computing ρ' using eigenvalue routines (see Sarmanov and Lancaster). For continuous random variables X,Y, the sup correlation is computable only in special cases where the joint p.d.f. admits a certain type of bivariate orthogonal expansion (see Lancaster).

An important dependence concept between two random variables is that of complete dependence, introduced by Lancaster (1963). A random variable Y is said to be completely dependent on a random variable X if there exists a function g such that

(1.1) $\quad \operatorname{Prob}[Y = g(X)] = 1.$

If Y is completely dependent on X and vice versa, then X and Y are said to be mutually completely dependent; in this case X and Y are perfectly predictable from each other. Observe that if X and Y are mutually complete dependent then $\rho'(X,Y) = 1$.

Kimeldorf and Sampson (1978) provided an example of random variables X and Y which were mutually completely dependent and yet were "almost" stochastically independent. To circumvent this difficulty, Kimeldorf and Sampson defined Y to

be <u>monotone increasing (decreasing) dependent</u> on X if (1.1) holds for a monotone increasing (decreasing) function g. Furthermore, motivated by trying to measure the degree of monotone dependence, they defined the <u>monotone correlation</u> between random variables X and Y by

$$(1.2) \qquad \rho^*(X,Y) = \sup \rho(f(X), g(Y)),$$

where the supremum is taken over all monotone functions f,g, for which $0 < \text{Var } f(X) < \infty$ and $0 < \text{Var } g(Y) < \infty$. The monotone correlation is a monotone scale invariant measure of association and the maximizing functions (assuming they exist) for (1.2) are the "best" monotone scalings for cross linear predictability of X and Y. (Monotone scalings are order-preserving assignments of numerical values to ordinal data.) Kimeldorf and Sampson evalated the monotone correlation in only the two special situations: (i) X and Y bivariate normal, in which case $\rho^* = |\rho|$, and (ii) X and Y independent, in which case $\rho^* = 0$.

The purposes of this paper are twofold: One is to derive new measures associated with the monotone correlation and to study their applicability. A second is to provide a computational procedure and computer program to evaluate the monotone correlation and these derived measures for the case when X and Y assume a finite number of values.[1] The approach is to find an equivalent nonlinear program and then employ a modification of the optimization algorithm of May (1979) to compute the maximizing values and the points at which they occur. In Section 2, we introduce the concepts of concordancy, discordancy, and isoscaling for measuring monotone association. The equivalent nonlinear programs are given in Section 3. The specific algorithm and the computer program, which we call MONCOR, are described in Section 4. A number of interesting applications

[1]The authors are in the process of examining procedures for data from certain continuous distributions.

and examples are considered in Section 5. In Section 6, we discuss how these methods might be used for scale reduction.

2. Concordancy, Discordancy, and Isoscaling

The concept of monotone correlation can be refined by measuring separately the strength of the relationship between X and Y in a positive direction and the strength of the relationship in a negative direction, that is, to measure separately the extent of concordancy and of discordancy between X and Y. These concepts are related to so-called measures of disagreement and measures of dissociation. If in (1.2), f and g are both restricted to be increasing[1] (or equivalently both decreasing), the resulting measure is called the concordant monotone correlation (CMC). When f is restricted to be increasing and g decreasing (or equivalently f decreasing and g increasing), we find it convenient to examine $-\sup \rho(f(X), g(Y))$, which in turn can be expressed as $-\sup \rho(f(X), -g(Y))$, where both f and g are increasing. This leads naturally to defining the discordant monotone correlation (DMC) by $\inf \rho(f(X), g(Y))$, where f and g are both restricted to be increasing.

The DMC and CMC have natural interpretations as measures of negative and positive association, respectively, for ordinal random variables. They also can be interpreted as providing bounds for the correlation between any arbitrary monotone scalings; specifically, for arbitrary increasing f and g,

$$(2.1) \qquad \text{DMC} \leq \rho(f(X), g(Y)) \leq \text{CMC}.$$

Suppose it is desired to impose numeric monotone scalings for a pair of new tests; if the CMC and DMC are close, then by (2.1) it makes little difference which monotone scales are used. Also if CMC = DMC = 0, then X and Y are inde-

[1] The terms increasing and decreasing are used non-strictly.

pendent random variables; however, it is possible for DMC < CMC = 0 and X and Y not to be independent. Further note that if X and Y are increasing monotone dependent then CMC = 1; and if X and Y are decreasing monotone dependent, then DMC = -1.

Sometimes the situation occurs when X and Y should have the same scaling. For example, X is a psychological test score pretreatment and Y is the score post-treatment on the same test. This leads to another extension of monotone correlation, which we refer to as isoscaling. If in (1.2) we restrict $f = g$, the resulting measure is called the isoconcordant monotone correlation (ICMC). Analogous to the DMC definition, the isodiscordant monotone correlation (IDMC) is given by $\inf \rho(f(X), g(Y))$, where $f = g$. Obviously isoscaling is not practically appropriate when X and Y have essentially different ranges of values.

If X and Y are exchangeable ordinal random variables it might be conjectured due to all the symmetries involved that ICMC = CMC (and IDMC = DMC). However, as is shown in Section 5, this is surprisingly not the case.

The actual functions that maximize the correlations (assuming they exist) are of importance in developing monotone scales. We refer to such functions generically as monotone variables with their specific interpretation depending on which monotone correlation measure is used in their derivation.

3. Program Formulation

The preceding extensions of the monotone correlation are applicable to all suitable pairs of random variables, continuous or discrete. We now focus on the case where X and Y jointly take on a finite number of values (a_i, b_j), $i = 1, \ldots, I$, $j = 1, \ldots, J$ and $\text{Prob}(X = a_i, Y = b_j) = p_{ij}$. Then

$$(3.1)\quad \mathrm{CMC} = \max \frac{\sum_{i=1}^{I}\sum_{j=1}^{J} f(a_i)p_{ij}g(b_j) - (\sum_{i=1}^{I} p_{i\cdot}f(a_i))(\sum_{j=1}^{J} p_{\cdot j}g(b_j))}{(\sum_{i=1}^{I} f^2(a_i)p_{i\cdot} - (\sum_{i=1}^{I} p_{i\cdot}f(a_i))^2)^{\frac{1}{2}}(\sum_{j=1}^{J} g^2(b_j)p_{\cdot j} - (\sum_{j=1}^{J} p_{\cdot j}g(b_j))^2)^{\frac{1}{2}}},$$

subject to f and g being increasing functions for which the denominator in (3.1) is non-zero and where $p_{i\cdot} = \sum_{j=1}^{J} p_{ij}$ and $p_{\cdot j} = \sum_{i=1}^{I} p_{ij}$. Denote the values $f(a_i)$ by x_i, $i = 1, \ldots, I$ and $g(b_j)$ by y_j, $j = 1, \ldots, J$, so that (3.1) can be reformulated as

$$(3.2)\qquad \mathrm{CMC} = \max \frac{\underset{\sim}{x}'\underset{\sim}{P}\underset{\sim}{y} - (\underset{\sim}{x}'\underset{\sim}{P}\underset{\sim}{e})(\underset{\sim}{y}'\underset{\sim}{P}'\underset{\sim}{e})}{(\Sigma x_i^2 p_{i\cdot} - (\underset{\sim}{x}'\underset{\sim}{P}\underset{\sim}{e})^2)^{\frac{1}{2}}(\Sigma y_j^2 p_{\cdot j} - (\underset{\sim}{y}'\underset{\sim}{P}'\underset{\sim}{e})^2)^{\frac{1}{2}}},$$

$$\text{subject to: } x_1 \le \cdots \le x_I$$
$$y_1 \le \cdots \le y_J$$
$$\underset{\sim}{x} \ne c_1\underset{\sim}{e},\ \underset{\sim}{y} \ne c_2\underset{\sim}{e}$$

where $\underset{\sim}{x} = (x_1, \ldots, x_I)'$, $\underset{\sim}{y} = (y_1, \ldots, y_J)'$, $\underset{\sim}{P} = \{p_{ij}\}$ and $\underset{\sim}{e} = (1, \ldots, 1)'$. Thus to compute the CMC all that is required is the matrix $\underset{\sim}{P}$ of probabilities. For instance, the values $a_1, \ldots, a_5$ could be the five-point scale strongly disagree, ..., strongly agree. The resultant monotone variable $\underset{\sim}{x}$ would then provide a numeric scale assigning x_1 to strongly disagree, ..., x_5 to strongly agree.

Analogous formulations of (3.2) can be given for ICMC, DMC, and IDMC. Again the ICMC and IDMC are not defined when $I \ne J$.

When reporting the monotone variables, we standardize them without loss of generality so that in (3.2), for example, $x_1 = y_1 = 0$ and $x_I = y_J = 1$.

Until this point, the CMC, etc., have been defined as population quantities. For data from finite discrete distributions, the joint probabilities can be estimated from the data viewed in ordinal contingency table form. Then the CMC can be evaluated based upon the estimated probabilities. In this situation, the

CMC can either be viewed as an estimate of the "true" CMC or be viewed as a measure of monotone association for the ordinal contingency table.

4. Optimization Approach and MONCOR Description

The nonlinear programming problem (3.2) involves the optimization of a nonlinear fractional form subject to linear constraints. Note that if it were not for the monotone constraints, (3.2) would be an eigenvalue problem. The objective function in (3.2) is not pseudoconcave. To see this, consider the simple case of evaluating the ICMC = $\max(\underset{\sim}{x}'\underset{\sim}{Px} - (\underset{\sim}{x}'\underset{\sim}{Pe})^2)/(\Sigma x_i^{\,2} p_{i\cdot} - (\underset{\sim}{x}'\underset{\sim}{Pe})^2)$ for a symmetric probability matrix $\underset{\sim}{P}$. While both numerator and denominator are continuously differentiable on the feasible region, and $(\Sigma x_i^{\,2} p_{i\cdot} - (\underset{\sim}{x}'\underset{\sim}{Pe})^2)$ is a positive convex function of $\underset{\sim}{x}$, $(\underset{\sim}{x}'\underset{\sim}{Px} - \underset{\sim}{x}'\underset{\sim}{Pe})^2$ would have to be nonnegative and concave for pseudoconcavity (see Avriel (1976)). This latter condition does not hold in general for symmetric $\underset{\sim}{P}$. Hence, in general, the CMC, and ICMC, DMC and IDMC, will involve the optimization of a function with local optima. Although much work is presently being done in the area of global optimization (see, for example, Dixon and Szego (1975), (1978)), we follow the standard procedure of using various starting points, computing the optima, and then choosing the best result based upon the different starting points.

Note that since correlation is unique in $\underset{\sim}{x}$ and $\underset{\sim}{y}$ only up to location and scale change, we could express (3.2) as

$$\text{(4.1)} \qquad \text{Maximize} \quad \underset{\sim}{x}'\underset{\sim}{Py} \,,$$

subject to $\sum_i x_i p_{i\cdot} = 0$, $\sum_j y_j p_{\cdot j} = 0$, $\sum_i x_i^{\,2} p_{i\cdot} = 1$, $\sum_j y_j^{\,2} p_{\cdot j} = 1$, $x_1 \le x_2 \le \cdots \le x_I$, and $y_1 \le y_2 \le \cdots \le y_J$.

The formulation of (4.1), because of its nonlinear constraints, is not a desirable formulation since complexity in the objective function is much easier to deal with than complexity in the constraints. The constraints $\underset{\sim}{x} \neq c_1\underset{\sim}{e}$ and $\underset{\sim}{y} \neq c_2\underset{\sim}{e}$ in (3.2) are not computationally implementable in continuous variables. However, without loss of generality, we eliminate those constraints by fixing x_1 and y_1 at zero and x_I and y_J at one.

Specifically, the computation of the CMC (DMC) involves optimizing a nonconcave (nonconvex) function in I + J - 4 independent variables subject to monotonicity constraints. (The ICMC and IDMC involve I - 2 independent variables.) Since $\underset{\sim}{P}$ is envisioned to be not much larger than 10 x 10, a modified Newton method was considered desirable because it should converge in a small number of iterations. QRMNEW (see May), an optimization method not requiring analytical derivatives, was employed because of its ease of adaptation and computational use.

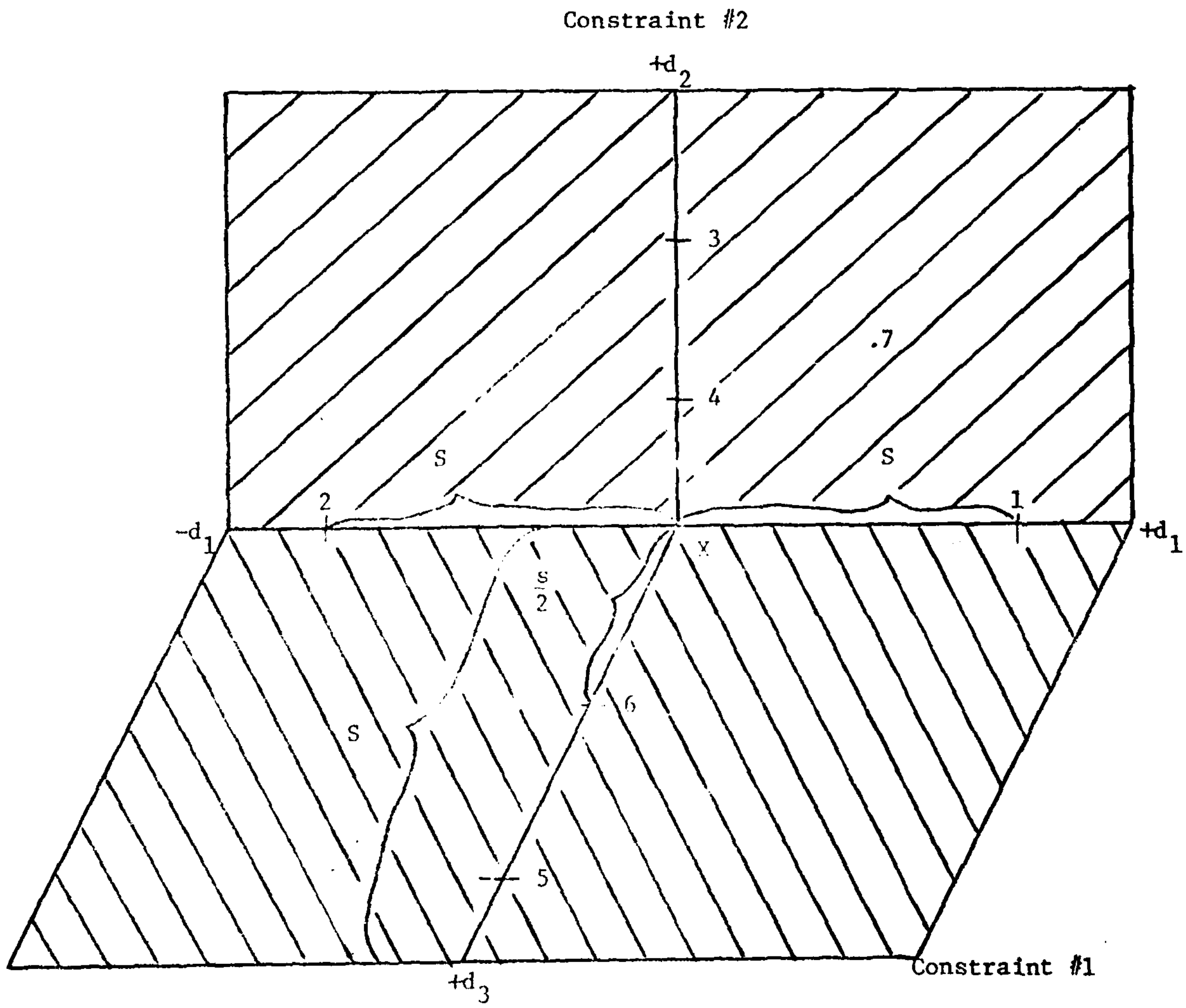
QRMNEW is a hybrid local variations-modified Newton method, using orthogonal (QR) matrix factorization to derive a representation for the locally feasible region. It has been proven globally convergent to a point satisfying both first and second order necessary optimality conditions, so that any solution generated is at least a local optimum. Superlinear and order 2 convergence rates can be established under somewhat stronger conditions. Denote by $\{(x,y)^k\}$ the iterative sequence of points generated by the algorithm. In general, because of the lack of pseudoconcavity (pseudoconvexity) for the CMC and ICMC (DMC and IDMC), an iterate $(x,y)^k$ will usually be in a region not locally concave (convex). The algorithm does have a rather sophisticated method for dealing with the indefinite projected matrix of second derivatives implied by the lack of local concavity (convexity).

While a complete mathematical description of QRMNEW is given by May, a general iteration is illustrated in Figure 4.1, to show the underlying logic. Given the current point x and a stepsize $s > 0$, the constraints, if any, that are satisfied exactly at x are used to generate, via orthogonal matrix factorization, a set of n coordinate directions. If u constraints are active, u directions lie in the manifold determined by those constraints, and the remaining n-u directions are determined by computing a generalized inverse and are orthogonal to that manifold. (If no constraints are active, the standard Cartesian coordinate system is used.) The objective function is then evaluated at 2 points along each of these coordinate axes. For example, in Figure 4.1, two constraints are active in R^3. Three directions are generated: d_1, which lies in the manifold, so that movement away from x in either $+d_1$ or $-d_1$ is feasible, d_2, where going along $+d_2$ leads to infeasibility, and d_3, which is analogous to d_2. The function is evaluated at points 1 through 6, yielding second order approximations to first and second partial directional derivatives along d_1, d_2, and d_3. Assume a maximum is being sought, e.g., computing the CMC or ICMC, and that the first derivatives along d_1, d_2, and d_3 are, respectively, positive, positive, and negative. Then the objective function cannot be increased by movement along d_3, so that it is dropped from consideration. The function is then evaluated at point 7, which is needed to approximate the second mixed partial directional derivative with respect to d_1 and d_2. A Newton-type search direction is computed and searched, and the algorithm moves to the best of the points found by the Newton search procedure and points 1 through 7.

MONCOR is an interactive package designed to analyze probability matrices, $\underset{\sim}{P}$, of dimension less than or equal to 20 x 20. The user may input a single starting point for an optimization run, or allow the program to generate its own multiple starting points. In both cases, the constraint set corresponding to the

Figure 4.1

An Iteration of QRMNEW



correlation measure requested is generated internally, and QRMNEW is used to compute an optimum. Additionally, two different strategies are employed in seeking an optimal solution. Numerical experience indicates that optimum values sometimes lie at monotone extreme points, i.e., points where all the $\tilde{x}$ and $\tilde{y}$ entries are either zero or one. This appears to be especially the case when computing the DMC or IDMC for a matrix with highly positive CMC, and vice-versa.

In fact, for certain cases, the optima for all four monotone correlation measures might be achieved only at such points. Additionally, because non-optimal monotone extreme points can be local optima (satisfying the Karush-Kuhn-Tucker second order necessary optimality conditions (see Kuhn and Tucker (1951))), QRMNEW starting from a random point might well be trapped by these local optima. Note that for an I x J matrix, there are only (I - 1)(J - 1) monotone extreme points to consider for the CMC and DMC ((n - 1) for ICMC and IDMC, assuming I = J = n). Hence, in order to avoid stopping at a local optima when the global optima is a monotone extreme point, MONCOR evaluates the correlation of all monotone extreme points. Moreover, MONCOR generates ten random monotone points, with coordinates selected on (0,1), using the DEC random number generator (see Payne, Rabung, and Bogyo (1969)), and calls QRMNEW to compute an optimum starting from each of them. The user may select to see only the final output, or an iteration-by-iteration output of the monotone correlations and monotone variables.

5. Applications

By means of the algorithm and the MONCOR program, we now compute the CMC, etc., for several insightful examples. Let (X,Y) be a discrete bivariate random vector taking values in a 6 x 6 lattice: $\{a_1, \ldots, a_6\} \times \{b_1, \ldots, b_6\}$. Further suppose $\text{Prob}(X = a_i) = 1/6$, for all i, and $\text{Prob}(Y = b_j) = 1/6$, for all j; i.e., X and Y have uniform marginals. If X and Y are monotone increasing dependent then $\underset{\sim}{P} = (1/6)\underset{\sim}{I}$, where $\underset{\sim}{P} = \{\text{Prob}(X = i, Y = j)\}$, and $\underset{\sim}{I}$ is the 6 x 6 identity matrix; if X and Y are monotone decreasing dependent then $\underset{\sim}{P} = (1/6)\underset{\sim}{I}^*$, where $\underset{\sim}{I}^* = \{\delta(i + j - 7)\}$, and $\delta(x)$ is 1 if x = 0, and is 0, otherwise. Now consider a one-parameter family of distributions indexed by θ, i.e., for a given θ, Prob(X = i, Y = j) is the (i,j)th element of $\underset{\sim}{P}_\theta$, where

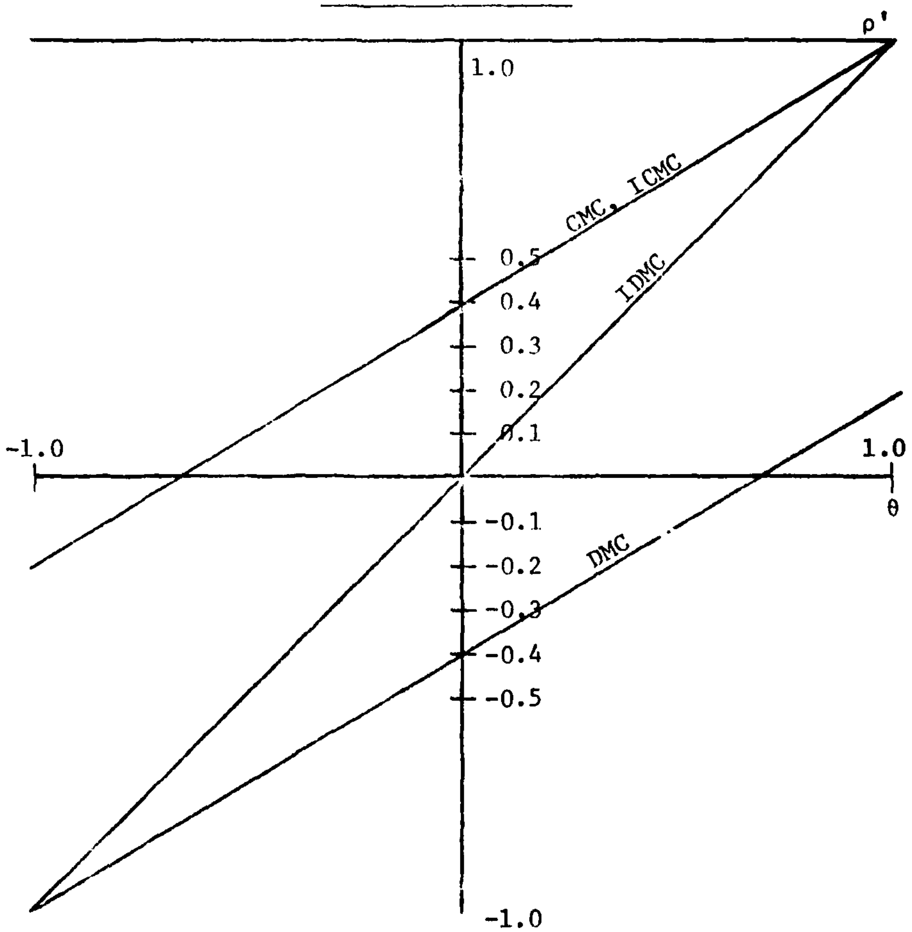
$$(5.1) \qquad \underset{\sim}{P}_\theta = \left(\frac{1+\theta}{2}\right)(1/6)\underset{\sim}{I} + \left(\frac{1-\theta}{2}\right)(1/6)\underset{\sim}{I}^*,$$

where $-1 \le \theta \le 1$. Note that X and Y still have uniform marginal distributions for all θ. For $\theta = 1(-1)$, $\underset{\sim}{P}_\theta$ corresponds to the most monotone increasing (decreasing) dependent case; and intermediate values of θ describe varying degrees of mixtures of the two dependent extremes. In Figure 5.1, we graph the values of the CMC, ICMC, DMC and IDMC as functions of θ for $\underset{\sim}{P}_\theta$ given by (5.1). (Moreover, because the support of X and Y is two disjunct pieces in the sense of Lancaster, it follows that the sup correlation ρ' is 1 for all θ in (5.1).)

Figure 5.1

CMC, DMC, ICMC, IDMC vs. θ

$\underset{\sim}{P}_\theta$ given by (5.1)



From Figure 5.1, it can be observed that the CMC, DMC, ICMC, and IDMC are all linear functions of θ. Moreover, the IDMC = θ. The CMC and ICMC coincide, and the CMC at θ is equal to the negative of the DMC evaluated at $-\theta$.

Now consider (X,Y) defined on a 3 x 3 lattice with

$$(5.2) \qquad \underset{\sim}{P} = \begin{bmatrix} 0 & 1/4 & 0 \\ 1/4 & 0 & 1/4 \\ 0 & 1/4 & 0 \end{bmatrix} ,$$

so that, for example, Prob($X = a_1$, $Y = b_2$) = 1/4. Note that $\underset{\sim}{P}$ is a symmetric probability matrix, so that X and Y are exchangeable random variables. It follows in this case by direct computation or by use of MONCOR, that the ICMC is 0, and the monotone variables for X and Y are (0, .5, 1)'. However, the CMC is 1/3, and the monotone variables for X and Y, respectively, are either (0, 1, 1)' and (0, 0, 1)' or (0, 0, 1)' and (0, 1, 1)'. Thus (5.2) provides an example of exchangeable random variables where ICMC $\neq$ CMC.

We now consider applying these monotone measures to an actual data example, taken from Bishop, Holland, and Fienberg (1975, p. 100), which in turn was adapted from Glass and Hall (1954, p. 183). These data are given in Table 5.1.

Because the same categories are used to measure father's and son's occupational status, it is appropriate to use isoscaling. The ICMC, IDMC and the associated monotone variables were computed by the MONCOR program based on the empirical probability matrix specified by Table 5.1. The values of the ICMC and IDMC as well as the monotone variables are presented in Table 5.2.

Table 5.1

British Mobility Data

(3,500 Father-Son Data Values)

Son's Occupational Status[1]

Father's Occupational Status[1]		S1	S2	S3	S4	S5
	S1	50	45	8	18	8
	S2	28	174	84	154	55
	S3	11	78	110	223	96
	S4	14	150	185	714	447
	S5	3	42	72	320	411

Table 5.2

ICMC, IDMC and Monotone Variables for

British Mobility Data

Measure	Value of Measure	Monotone Variable Values				
ICMC	.496	0.	.627	.842	.923	1.0
IDMC	.242	0.	0.	0.	0.	1.0

The analogous version of (2.1) for isoscaling, namely IDMC $\leq \rho[f(X), g(Y)] \leq$ ICMC, shows that regardless of the assignment of numerical values to the five ordinal categories, the resultant correlation is between .242 and .496.

[1] Status S1 is professional, and high administrative; status S2 is managerial, executive, and higher grade supervisory; status S3 is lower grade supervisory; status S4 is skilled manual; and status S5 is semi-skilled and unskilled manual.

6. Scaling

One important use of monotone variable theory is the ability to develop meaningful scales for ordinal variables. For example, suppose the five-point scale response to some question is elicited pre- and post- some experimental intervention. Through the use of the ICMC, we can provide a numerical scale for this five point response; this numerical scale has the property that among all possible such ordinal scalings, the post-response for this scaling is most linearly predictable from the pre-response. In Table 5.2 the row corresponding to ICMC provides this scaling for the occupational status variable based on the British mobility data. Specifically, the numerical values for S1, S2, S3, S4, and S5 are 0., .627, .842, .923, and 1.0, respectively.

Often, the number of distinct values for the numerically scaled variables is substantially less than the number of values for the original ordinal variables. This reduction occurs when the optimizing f,g in (1.2) are not one-to-one functions. To illustrate this phenomenum, we consider the following example. A 10×10 matrix is generated where each entry is a randomly generated number on (0,1), each generated independently of the other entries. In order to generate a "slightly" positive dependent distribution, the constant 2 was added to each diagonal and the entire matrix scaled so as to add to one. The resultant matrix is given in Table 6.1.

The CMC for the matrix in Table 6.1 is 0.443, and the monotone variables for a_1, ... a_{10}, and b_1, ..., b_{10}, are, respectively, (.000, .461, .461, .461, .872, .872, .872, .872, .873, 1.000)' and (.000, .537, .541, .541, .842, .842, .842, .842, .842, 1.000)'. Note that while the original variables each had 10 separate values, there are only five distinct monotonely scaled values for X and five for Y. While this scale reduction phenomenum is based upon empirical observation, it is clear that it has great potential value in deriving simplified scales for large data sets.

Table 6.1

Random 10 x 10 Probability Matrix

X \ Y	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9	b_{10}
a_1	.0331	.0111	.0092	.0049	.0016	.0028	.0009	.0108	.0096	.0007
a_2	.0101	.0361	.0057	.0081	.0133	.0062	.0121	.0066	.0003	.0020
a_3	.0102	.0059	.0347	.0027	.0055	.0020	.0104	.0046	.0069	.0056
a_4	.0144	.0018	.0065	.0342	.0006	.0071	.0055	.0066	.0084	.0113
a_5	.0006	.0016	.0087	.0132	.0435	.0061	.0100	.0046	.0044	.0053
a_6	.0022	.0035	.0151	.0015	.0056	.0427	.0062	.0035	.0089	.0125
a_7	.0002	.0084	.0026	.0020	.0005	.0086	.0387	.0007	.0034	.0111
a_8	.0084	.0100	.0079	.0036	.0100	.0128	.0044	.0303	.0121	.0065
a_9	.0028	.0079	.0141	.0008	.0133	.0077	.0064	.0139	.0402	.0068
a_{10}	.0009	.0149	.0042	.0108	.0022	.0144	.0130	.0151	.0146	.0438

7. Program Availability

The MONCOR program, written in FORTRAN, is available for distribution. For specific details contact Professor Jerrold May, Graduate School of Business, University of Pittsburgh, Pittsburgh, PA 15260.

REFERENCES

1. Avriel, M., (1976). Nonlinear Programming: Analysis and Methods, Prentice-Hall, Inc., Englewood Cliffs, NJ.

2. Bishop, Y.M.M., Fienberg, S.E., and Holland, P.W., (1975). Discrete Multivariate Analysis: Theory and Practice, MIT Press, Cambridge.

3. Dixon, L.C.W., and Szego, G.P., (1975). Towards Global Optimisation, North-Holland Publishing Company, New York.

4. Dixon, L.C.W., and Szego, G.P., (1978). Towards Global Optimisation, 2, North-Holland Publishing Company, New York.

5. Gebelein, H., (1941). "Das Statistiche Problem der Korrelation als Variations und Eigenwert problem und sein Zusammenhang mit der Ausgleichungsrechnung." Z. Angew. Math. Mech. 21, 364-379.

6. Glass, D.V. and Hall, J.R., (1954). "Social mobility in Britain: A study of inter-generation changes in Status." In Social Mobility in Britain. Ed. by D.V. Glass. Routledge and Kegan Paul Ltd., London.

7. Kimeldorf, G. and Sampson, A.R., (1978). "Monotone dependence." Annals of Statistics 6, 895-903.

8. Kuhn, H.W., and A.W. Tucker, (1951). "Nonlinear Programming," in J. Neyman (ed.), Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, pp. 481-492. University of California Press, Berkeley.

9. Lancaster, H.O., (1963). "Correlation and complete dependence of random variables." Ann. Math. Statist. 34, 1315-1321.

10. Lancaster, H.O., (1969). The Chi-Squared Distribution. Wiley, New York.

11. May, J.H., (1979). "Solving Nonlinear Programs Without Using Analytic Derivatives," Operations Research, 27, 457-484.

12. Payne, W.H., Rabung, J.R., and Bogyo, T.P., (1969). "Coding the Lehmer Pseudo-random Number Generator," Communications of the ACM, 12, 85-86.

13. Rényi, A., (1959). "On measures of dependence," Acta. Math. Acad. Sci. Hungar. 10, 441-451.

14. Sarmanov, O.V., (1958a). "The maximal correlation coefficient (symmetric case)." Dokl. Akad. Nank. SSSR 120, 715-718. (English Translation in Sel. Transl. Math. Statist. Probability 4, 271-275.)

15. Sarmanov, O.V., (1958b). "The maximal correlation coefficient (non-symmetric case)." Dokl. Akad. Nank. SSSR 121, 52-55. (English Translation in Sel. Transl. Math. Statist. Probability 4, 207-210.)

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 80-1344	2. GOVT ACCESSION NO. AD-A093816	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Concordant and Discordant Monotone Correlations and Their Evaluation by Nonlinear Optimization		5. TYPE OF REPORT & PERIOD COVERED INTERIM
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) George Kimeldorf, Jerrold H. May and Allan R. Sampson		8. CONTRACT OR GRANT NUMBER(s) F49620-79-C-0161
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Pittsburgh Department of Mathematics & Statistics Pittsburgh, PA. 15260		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research /NP Bolling Air Force Base Washington, D. C. 20332		12. REPORT DATE October 1980
		13. NUMBER OF PAGES 17
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Monotone correlation, concordant monotone correlation, discordant monotone correlation, isoscaling, ordinal contingency tables, nonlinear optimization		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) This paper presents four new statistical measures of monotone relationship derived from the concept of monotone correlation. A nonlinear optimization algorithm is employed to evaluate these new measures, as well as the monotone correlation, for ordinal contingency tables. A computer program to implement the algorithm is developed, and is applied to several insightful examples to provide further understanding of the usefulness of these measures.		

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)